



This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Laaki, Heikki; Miche, Yoan; Tammi, Kari

Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks : Application of Remote Surgery

Published in: IEEE Access

DOI:

10.1109/ACCESS.2019.2897018

Published: 01/02/2019

Document Version
Publisher's PDF, also known as Version of record

Please cite the original version:

Laaki, H., Miche, Y., & Tammi, K. (2019). Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery. *IEEE Access*, 7, 20325-20336. Article 8632888. https://doi.org/10.1109/ACCESS.2019.2897018

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.



Received December 10, 2018, accepted January 11, 2019, date of publication February 1, 2019, date of current version February 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2897018

Prototyping a Digital Twin for Real Time Remote Control Over Mobile Networks: Application of Remote Surgery

HEIKKI LAAKI¹⁰¹, YOAN MICHE², AND KARI TAMMI¹⁰¹, (Member, IEEE)

Department of Mechanical Engineering, Aalto University, 02610 Espoo, Finland

²Nokia Bell Labs, 02610 Espoo, Finland

Corresponding author: Heikki Laaki (heikki.laaki@aalto.fi)

This work was supported by Tekes as part of the DIMECC Cyber Trust Program.

ABSTRACT The concept of digital twin (DT) has emerged to enable the benefits of future paradigms such as the industrial Internet of Things and Industry 4.0. The idea is to bring every data source and control interface description related to a product or process available through a single interface, for auto-discovery and automated communication establishment. However, designing the architecture of a DT to serve every future application is an ambitious task. Therefore, the prototyping systems for specific applications are required to design the DT incrementally. We developed a novel DT prototype to analyze the requirements of communication in a mission-critical application such as mobile networks supported remote surgery. Such operations require low latency and high levels of security and reliability and therefore are a perfect subject for analyzing DT communication and cybersecurity. The system comprised of a robotic arm and HTC vive virtual reality (VR) system connected over a 4G mobile network. More than 70 test users were employed to assess the system. To address the cybersecurity of the system, we incorporated a network manipulation module to test the effect of network outages and attacks; we studied state of the art practices and their utilization within DTs. The capability of the system for actual remote surgery is limited by capabilities of the VR system and insufficient feedback from the robot. However, simulations and research of remote surgeries could be conducted with the system. As a result, we propose ideas for communication establishment and necessary cybersecurity technologies that will help in developing the DT architecture. Furthermore, we concluded that developing the DT requires cross-disciplinary development in several different engineering fields. Each field makes use of its own tools and methods, which do not always fit together perfectly. This is a potentially major obstacle in the realization of Industry 4.0 and similar concepts.

INDEX TERMS Digital twin, virtual reality, robot control, mobile networks, network security.

I. INTRODUCTION

Industrial Internet and Internet of Things (IoT) have brought up a capability to measure and observe various phenomena. Data available from those has in turn created an opportunity to develop system modelling capabilities such as Digital Twins (DT). The concept of DT has emerged to make the physical things and related data sources accessible for software and users on digital platforms [2]. Terms used for similar or partially overlapping concepts include digital counterpart, virtual twin, virtual object, product agent, and avatar [3]. DT is argued to consist of three parts: physical

The associate editor coordinating the review of this manuscript and approving it for publication was Okyay Kaynak.

product, digital product and the linkage between the physical and digital products [4]. DT enables real time communication with the physical twin in both directions. Interested entities, such as processes or systems, can read data from the physical product via a DT, analyze the data and execute simulations. Vice versa, actuation commands, control signals and other data can be sent to the physical twin with the help of a DT. The DT contains a structured description of available communication and control interfaces of the physical asset. These interfaces can be accessed and harnessed for application specific needs by analyzing the DT and without accessing and examining the physical twin at the site. Apart from the communication aspect, a DT will serve as container of all existing data about the physical twin. The actual data, such



as CAD model files, sales documentation and measurement data, technically exist in several different software and systems. The role of DT is to bring these various data stores accessible via a single interface.

Furthermore, just like a physical object is always tied to specific location and time, a DT needs to incorporate location and time aspects despite of it's digital nature. These fundamental features are the core of a DT. The ultimate value proposition of a DT is to enable communication between devices and systems in real time and, more importantly, automatically. Without DT, setting up complex systems, such as our remote surgery system, would require specialists and experts for each product, device and software piece incorporated. By analyzing the DTs of the incorporated components, developers and engineers can determine, design and create the required interfaces, integrations and communication links without specific expertise of each component. Eventually, the devices might be able to discover and understand each other without any human engineer in between. This sort of auto-discovery and auto-established communication with the help of DTs could eventually make IoT more scalable for applications currently unimaginable. The underlying problems, beside the obvious technical one that such protocols and networking architecture are currently only under development [2], include authentication of each device and user, safety and security, data ownership issues, reliability, robustness and acceptable delays, among others. Solving these problems with DTs will enable the Industrial Internet and IoT to really add value to current systems and processes.

We developed a novel prototype system to investigate what kind of communication link is required in a mission critical application. As a case application, we chose the specific task of remote surgery. Such operations require near zero latency and high levels of security and reliability [5] and therefore are a perfect subject for analysis in the context of DTs involved in achieving the required features and functionality. Defining proper DT architecture that has potential to succeed in every possible future application is an immensely ambitious task. This kind of analysis via specific prototype applications provides the necessary building blocks to eventually be able to determine the structure and architecture of DTs. Moreover, we could determine if mobile network supported medical operations are feasible. We collected experimental results about the functionality of the DT, different feedback technologies and usability of the system in test sessions with test users. Cybersecurity was studied with implementations regarding network security and user authentication and analyzing state of the art technologies and their consequences in development of DT concept. Based on the experiences gained during the development phase and experiments, we present ideas on how a DT should be designed to succeed in a mission critical application with special emphasis on the cybersecurity issues.

In many cases, the subject of a medical operation is in critical state and delays might result in further traumas. Additionally, it might be expensive, inconvenient or impossible for the patient and the medical personnel to travel to a mutual location. In these cases, remote operations could prove useful and might be possible with modern technologies such as Virtual Reality (VR), robotics and communication via a DT over mobile networks. VR supported applications provide great feedback and a significant improvement compared to more traditional video link and instruction passing. The first commercial VR applications, such as Forte Technology's VFX1, were launched in 1994 [6]. Only recently have the required technologies advanced enough to bring the VR devices to really usable level [7]. The best devices are able to generate virtual worlds that appear and feel almost like reality [8]. Moreover, monitoring various data points in real time from the user has become possible with smart phones and without extensive purpose-built sensor setups. This data can be used in augmenting virtual realities with real life elements [9].

Current state of the art technology in robotic surgery includes Da Vinci system. It enables prostate cancer surgeries to be performed with a robot inside a patient. A VR controller device is situated close to the patient. A doctor has a VR view inside the patient and maneuvers a robot with special controller interface [10], [11]. Furthermore, some surgeries have been performed remotely. One example is a cholecystectomy (surgical removal of gallbladder) performed over a distance of 6230 km between the surgeon and the patient in 2001 [12]. However, the connection used was specially set up for the particular operation, which means that using similar technology requires these connections set up beforehand in order to perform operations quickly on demand. The problems in remote control of robots in surgical applications include sufficient feedback for the surgeon, timely co-operation and communication with the remote and local teams, calibration of the control system with the patient and reliability, latencies and cybersecurity of the connection, among others. Using digital technology (in contrast to analog technology) enables digital concepts such as DT and AI in solving these problems and utilizing existing digital networks such as Internet and mobile networks. Remote surgeries using mobile networks has not been attempted to date. Such operations require exceptionally safe and trusted hardware, software and communication systems. With the onset of 5G mobile network technology, robotic telesurgery over 5G has become a research hotspot to resolve the various issues such as required human-machine interfaces, security, privacy and reliability [13]. In addition to contribution in development of DT concept, our prototype system brings remote surgeries over mobile networks closer to reality by providing solutions to the above mentioned problems.

II. TEST SETUP

While actual surgeries are complex and include several systems and personnel operating simultaneously, we simplified the problem into the bare essentials of using a set of instruments to operate a patient. By using a VR system to control an industrial robotic arm, we could achieve this core function.





FIGURE 1. A robotic arm was used to move a set of instruments to simulate a remote surgery of a dummy patient.

The simplified (physical) surgery environment is illustrated in figure 1.

VR could refer to plenty of distinct concepts, such as the virtual world described in The Matrix film series [14] or video games that are built around a virtual world in which the player can complete missions. With VR, we refer to a technology where the user is completely immersed into a virtual world with a head mounted display (HMD). The technology utilizes various sensors to create an immersive experience where the user is actually part of the VR and is able to explore and interact with the VR world. Similar sensors can be attached to other physical objects to bring them into the VR world as well. This sort of VR technology is used for variety of purposes such as game industry [15], architecture visualization [16], art creation [17], engineering [18] and movie industry [19]. These VR applications simulate real world or some real world elements depending on the requirements of each application.

The developed system required both the possibility to display and manipulate a virtual world. HTC Vive system [20] fulfilled these requirements plus had the highest rated reviews on the internet [21]. Other possible options included Oculus Rift with its Oculus Touch controllers [22]. The HTC Vive has a HMD unit and two handheld controllers. The controllers have an extensive interface for different actions to manipulate the virtual world and a haptic feedback function. The system utilizes an infrared laser grid to track the movement of the HMD and controllers. Inertial measurement units are used to improve the location tracking within the laser grid. The specification of the Vive system is presented in table 1.

Industrial robotic arms are generally used in mass manufacturing, which includes highly repetitive work and requires accuracy for extended periods of time. The shift from human workforce to robots requires substantial investment which

TABLE 1. HTC Vive specification according to Kreylos [23] and HTC [20].

Display refresh rate	90Hz
Field of view	110°
HMD location refresh rate	225Hz
Controller location refresh rate	250Hz
Tracking residual noise range	0.3mm
Tracking error (RMS)	2mm

can be more easily compensated in such manufacturing processes. Typical tasks for a robot are monotonous, difficult, dangerous or otherwise unpleasant for a human to complete. [24] For example, applying an even layer of paint on a large area can be difficult and include harmful working conditions for a human. However, such a task can be programmed to a robot [25].

Recently, robotic arms have started to share the working space with human workers. Traditionally robots have been working in separate spaces, but with advanced safety measures the collaboration with humans has become feasible. These collaborative robots are specifically called cobots. Robots must comply with EN ISO 10218-1:2011 safety standard and ISO/TS 15066:2016 technical specification, which specify safety measures for cobots in industrial applications [26], [27]. The safety requirements can be met for example with a collision detection system. However, each environment and application is unique and safety has to be evaluated case by case.

The typical advantages of industrial robotic arms were not especially utilized in this study. The most important feature utilized was the robotic arm's ability to replicate human hand movement, which requires six degrees of freedom. An industrial robotic arm with six joints fulfills this requirement [24] and thus suits well as a remote replicator of the user's

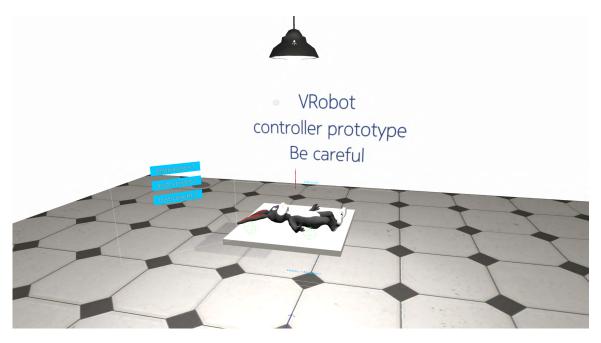


FIGURE 2. The Virtual Reality scene included a simplified surgery environment with a virtual dummy patient.

hand movements. As described, a robot operating in the same working space with humans requires special safety measures to avoid accidents. A prototype system is not an industrial application and does not necessarily require such measures. Nevertheless, the prototype system will be used for research and possibly for demonstrations with personnel unfamiliar with the system. Moreover, humans are definitely present in the same working space in medical tasks and hence the arm needed to be a cobot.

Universal Robots UR3 met the requirements specified above. Furthermore, the UR3 had an extensive scripting language for programming the robot, which was seen potentially helpful in building the prototype [28]. Moreover, the UR3 has an interface for different end effectors which helps in modifying the prototype system for different applications. The reach of the arm was slightly limited, but sufficient for research purposes and demonstrations. The specification of the UR3 is presented in table 2. [29]

TABLE 2. Universal Robots UR3 specification [30]. Collaboration operation refers to the concept of human workers and a robot working in the same working space simultaneously.

Repeatability	0.1mm
Payload	3kg
Reach	500mm
Degrees of freedom	6 rotating joints
Collaboration operation	15 advanced adjustable safety functions

Comparing the specification of the HTC Vive system and the UR3 arm, as specified in the tables 1 and 2 above, the former seems to be the limiting factor in terms of accurately replicating the hand movements of the user. However, based on our tests with the system, the achieved

accuracy is ultimately limited by the accuracy of the human hand. Consequently, we scaled the movements in VR down by a factor of 0.5 as described in section II-B, which reduces the human hand accuracy. Furthermore, this reduces the effect of HTC Vive inaccuracy in half, i.e. the resulting tracking error is 1mm (RMS) based on the measurements made by Kreylos [23]. The robotic arm controller is running at a frequency of 125 Hz and needs to be given instructions once each run [28]. As seen from the specification, the refresh rate of the HTC Vive system is considerably greater. Consequently, the achieved update rate of the system is limited to 125 Hz in optimal network conditions. Once the update rate drops below 125 Hz, the robot performs an emergency stop.

A. VR SCENE AND ROBOT SITE

The VR hardware requires a software framework to create any VR experience. Unity is a platform for developing 3D games and it is widely used in creating VR experiences and games. Unity handles all background processes, graphics processing and provides a physics engine for creating realistic virtual worlds. The developer only needs to program the game logic and design the virtual world scene [31]. The hardware manufacturer HTC and gaming software company Valve have collaborated to create a software development kit (SDK) called OpenVR, which provides software development tools for creating VR content. The SDK contains an interface for translating the tracking data from HTC Vive to a format that Unity is able to understand as well as using the haptic feedback and other controller functions [32].

Utilizing Unity and OpenVR, we created a VR scene that resembles loosely a medical setting for the prototype system. In the scene, a virtual dummy patient was located on top of



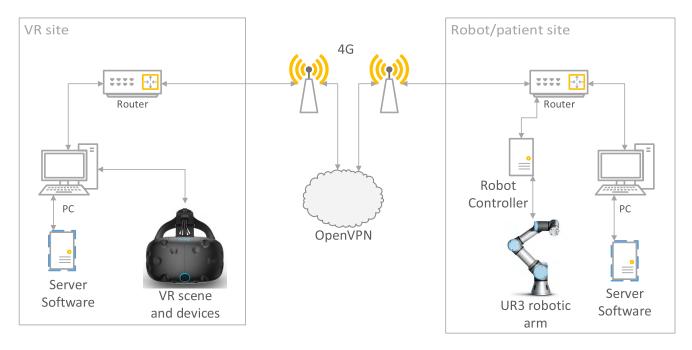


FIGURE 3. Network setup diagram of the prototype system.

a virtual table to mirror a physical dummy patient (stuffed bunny), which was on a table in real world. Their locations were calibrated to match. Moreover, the virtual dummy patient was scaled up, so that the VR movements could be scaled down for improved accuracy. A virtual representation of a medical instrument was floating above the patient. This instrument was used as the interface to control the robot intuitively. A user needs to pick up this virtual instrument and move it around however required and the robotic arm replicates that movement in real life. Three buttons were located close to the virtual operation table for the user to request different instruments. An illustration of the scene is shown in figure 2.

For the UR3 arm to be able to pick up instruments, we attached a simple electric magnet to the end of the arm. Using the UR3's built-in end effector interface we were able to switch the magnet on and off as required automatically or under user command. A teaching sequence was created with which the user could teach the robot how and where to pick up required instruments, so that the instruments could be switched automatically without human intervention at the robot site.

The system was separated at two sites, the VR site and robot site. Both sites were running a server computer that handled the networking and most of the data processing. The VR site server computer was also running the VR scene. Furthermore, a robot controller handled the processes that were closely related to the robotic arm. The two sites were interconnected with a 4G mobile network connection. The connection was established through a OpenVPN server which was running on a cloud server in Helsinki, capital of Finland. A schematic of the network setup is shown in figure 3.

B. DIGITAL TWIN

Once the hardware and network were set up, we started building the communication within the system which finally completes the DT of the system. It includes physical products, digital products and the connectivity in between, as defined in e.g. [4].

When using the system, a user picked up the virtual instrument using a hand held controller. The VR hardware tracked the movement of the controller and delivered the location and rotation data to the VR scene. The data was interpreted in the OpenVR software layer and the result was location and rotation data relative to the VR scene. Using that data, the system was able to tell the robotic arm what pose it should reach for. However, the coordinate systems of the VR and the robotic arm were different and the VR software does not understand robot poses. Moreover, the system had other limitations that had to be addressed. Consequently, several translations and conversions had to be applied before the robotic arm could utilize the data.

The Unity based VR world has a left handed coordinate system with X-axis to the right, Y-axis up and Z-axis forward with the location data stored in a quaternion format. On the contrary, the UR3 controller uses a right handed coordinate system with X-axis to the right, Y-axis to forward and Z-axis upward with the data stored in an axis-angle vector [28]. First, the rotation data had to be translated from quaternion format to an axis-angle vector format with equation

$$\begin{cases} \alpha = 2 * acos(qw) \\ x_{vr} = qx/\sqrt{1 - qw * qw} \\ y_{vr} = qy/\sqrt{1 - qw * qw} \\ z_{vr} = qz/\sqrt{1 - qw * qw}, \end{cases}$$
(1)



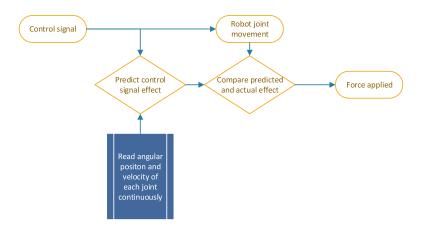


FIGURE 4. Illustration of the robot controller force estimation process.

where α is the angle in the axis-angle format. x_{vr} , y_{vr} and z_{vr} are the rotation axis components in the VR scene coordinate system. qx, qy, qz and qw are the components of the rotation quaternion.

The UR3 controller stores the angle component as the length of the axis vector [28]. Thus, the resulting axis had to be normalized and multiplied by the angle α from equation 1, which is done in equations

$$l = \sqrt{x_{vr}^2 + y_{vr}^2 + z_{vr}^2} \tag{2}$$

and

$$\begin{cases} x_{vr} \leftarrow x_{vr} * \alpha/l \\ y_{vr} \leftarrow y_{vr} * \alpha/l \\ z_{vr} \leftarrow z_{vr} * \alpha/l, \end{cases}$$
 (3)

where l is the length of the original vector.

Finally, to address the different coordinate systems, the basis had to be changed. Moreover, the robotic arm is located in the middle of its coordinate system and unable to operate close to that area because it would collide itself. Thus, the center of the coordinate system had to be offset. For actual usability, a favorable position for the robotic arm was behind the operation table. Thereby free operating space for assisting work and observing the work could be achieved. These requirements were fulfilled by adding an offset. The resulting equation is

$$\begin{cases} x_r = -x_{vr} \\ y_r = -z_{vr} - s_{offset} \\ z_r = -y_{vr}, \end{cases}$$
(4)

where x_r , y_r and z_r are the components of the axis-angle vector in the robot coordinate system with the angle as the length of the vector. s_{offset} is the offset added.

The robotic arm had a reach of 0.5m measured from the middle of the robot base. The constraints can be formulated more specifically as follows.

 The robot cannot operate outside of the spherical reach area at all.

- The robot operation is limited close to the edge of the spherical reach area.
- The robot cannot operate close to the robot base.
- The robot cannot operate below the operation table.

To address these constraints, a so called safety box was created. The system simply ignored any coordinates outside of the safety box and replaced them with values within the limits. The maximum usable dimensions with the UR3 robot were width 0.4m, depth 0.3m and height 0.25m. The movements in the VR scene were scaled down with a factor of 0.5, because the virtual dummy patient was scaled up. Combining the calculations and the safety box, a conversion between VR and robotic arm systems was achieved and a prototype DT was created.

Inside each site, the network connection setup utilized Transmission Control Protocol (TCP). The protocol incorporates means for ensuring reliable, ordered and error-checked delivery of packets [33]. Such features were desirable to accomplish the reliability required for our mission critical application. However, between the sites it was more beneficial to reach rapid delivery of the packets over the mobile network connection. This was achieved by using User Datagram Protocol (UDP) [33]. With UDP, we were able to broadcast an excess of packets, but we had to develop our own error-checking to ensure proper functionality. A simple order check proved to be sufficient.

Beginning from the VR scene program, the program first gets the location data from the VR scene. The safety box and conversions are applied before sending the data to a server software among with some other data. The VR site server software adds a running counter value to the message and broadcasts it to the robot site server over the 4G connection. The traffic is routed through the attack simulation module before reaching the robot site server software. The robot site server rejects messages that are received in wrong order and forwards the accepted data to the UR3 robot controller.

The UR3 programming language has built-in functions for moving the arm to a desired pose. A pose is a combination of exact position and orientation of the robot tool head.



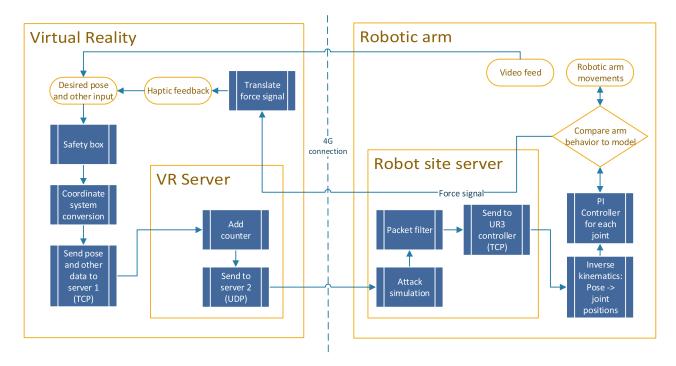


FIGURE 5. Illustration of the control loop data flow.

However, a dynamically updating goal pose is not allowed with these built-in functions. Using these functions resulted in a trembling movement as the robot stopped for a short moment every time it reached a pose and read the next desired pose. Ravn et al. had noticed that a dynamically updating desired pose can be achieved only by controlling the joint servo motors directly [34]. This was accomplished with the built-in inverse kinematics function and implementing a PI controller for each joint, which made the arm follow the control signal smoothly. Moreover, the controller continuously evaluated if the robot behaved as expected based on a simulation model and produced a force signal as a result. The force estimation process is illustrated in figure 4.

The force signal is returned to the robot server which is then forwarded back to the VR server and eventually to the VR scene. The signal is used to drive the haptic feedback device included in the handheld controllers of the HTC Vive. As a result, the system incorporates a force feedback for the user. Whenever the robotic arm experiences any force in any direction, the handheld controller provides a feedback through the haptic vibrator device. This force feedback proved to be somewhat insufficient, as noted in section III.

In addition to the robot location control, the DT program also delivers requests to initiate a tool change process, control the end effector or block the robot operation in case of compromised security. The control loop data flow is illustrated in figure 5. A more detailed description of the control loop program, data delivery and medical applications can be found in a thesis which was written about the control system [35].

We wanted be able to test how the system behaves under sub-optimal network conditions, such as under an attack. Hence, we added a network traffic manipulation module on the robot site server. With the module we simulated different network problems such as delays and attacks. Moreover, we wanted to test how certain mitigations work in our mission critical application context. We implemented the neural network based Denial of Service countermeasure which was developed by Kalliola et al. [36]. This method monitors network traffic metadata (network flows features), creates clusters of similarly behaving nodes and subnetworks, and trains numerous fast neural networks to recognize the features of typical traffic for each of the established clusters. With this model, normal network traffic (i.e. the traffic that the model has trained to recognize) is then prioritized, while the non-normal, i.e. potentially malicious, traffic only uses the potentially remaining network bandwidth, without possibility for over saturation. We refer the reader to the original publication for more details on the approach, which also treats other types of network anomaly events. This sort of mitigation should work well in mission critical applications, such as remote surgery, against attacks that could not be classified and detected reliably in advance [36].

III. EXPERIMENTAL RESULTS

The robot was able to replicate human hand movement accurately with the developed control loop setup. In case of really fast movement, the arm lagged behind the desired pose. The required speed for each joint simply exceeded the maximum speed of some of the joints. Moreover, as each joint had different speed limitations, the resulting trajectory was slightly curved and some sharp changes in direction became rounded when moving rapidly. Nevertheless, the behavior



was acceptable in most situations and a viable operation level was reached. A video demonstration of the control loop functionality is available at [37]. The described features can be distinguished in the video.

We were able to maneuver the robot delicately enough to pet the stuffed bunny gently and potentially perform some actual medical task. To further test the usability and behavior of the system, we demonstrated the system with more than 70 test users in several (identical) settings. The robotic arm site and the VR site were located roughly 10km apart in one of the test sessions and a separate video stream was provided for the audience and us to see the happenings at the other location. Smooth remote operation of the system can be seen in a video created from this session. The video is available at [38]. The robotic arm and VR site were located a few meters apart in rest of the test sessions, while the communication was still routed through the remote VPN server and the resulting network delays remained the same.

Based on the user test sessions, we divided the participants into three groups. Group 1) consisted of persons already familiar with VR technologies. The individuals within the group were able to control the robot easily when immersed in VR. When real life features, such as visual feedback, were mixed into the scene, it became harder to adapt, but the individuals quickly reached a similar level of accuracy to the system developers. Group 2) included persons testing this kind of VR technology for the first time. These individuals needed some time to familiarize with the VR system before they were able to effectively control the robot. Persons in group 3) were not able to adapt to the VR system at all, because they either failed to orient themselves in VR or Virtual Reality Induced Symptoms and Effects (VRISE), such as vertigo and loss of situational awareness. Root causes for VRISE include poor frame rate and sensory conflicts between the visual and the vestibular senses [39]. VRISE were present especially while using the visual feedback. These individuals could not control the robot accurately at all, but these effects usually faded after using the system for a while. Approximately 15% of people belonged to group 1), 80% to group 2) and 5% to group 3).

Force feedback plays a significant role in robotic surgeries as concluded e.g. in [40]. Consequently, we developed a force feedback loop into our system. The haptic device inside the hand held controller received a force feedback signal from the robot and vibrated if force was applied. The vibration intensified in relation to the magnitude of the force. The force feedback provided some help in sensing the contact between the instrument and the subject patient. However, we noted that the force signal is quite ambiguous with small forces. Significant improvement would be needed for actually usable force feedback.

To further improve the usability of the system and fully exploit the capabilities of VR technology, we developed a visual feedback system. VR is a concept where the user is immersed into a virtual world and real world elements are mostly removed. Augmented reality (AR) and mixed reality (MR) are similar concepts with the difference in real world

elements included. We created a mixed reality experience by bringing a video feed from the real life into the VR scene. We equipped the robotic arm with a video camera. With this setup, the users were able to look around in the physical world

Two different MR versions were tested. The first version made the camera follow the movement of the hand held controller. The video feed was displayed on a floating screen in the VR, which also followed the hand movement and felt like an imaging instrument. The second version followed the user's head, which let the user look around in the physical world in a natural way. In this case, the video feed was displayed directly in front of the user's eyes, which created a feeling of actually standing where the robotic arm stands. These visual feedback modes helped the users a lot more compared to the force feedback. However, the quality of the video feed was inadequate.

Under optimal network conditions and without experiencing any attacks, the system functioned and behaved smoothly and nearly without delays. Some tests have concluded that lag times between 30 to 150 ms are completely undetectable by surgeons [41]. The delays of our system were not precisely measured, but remained undetectable and below 150ms. Furthermore, we wanted to see how the system behaves under sub-optimal network conditions and even while under attack. Hence, we used the network traffic manipulation module to simulate a Distributed Denial of Service (DDoS) attack. A DDoS attack floods the communication link with malicious traffic and consequently interferes with the delivery of normal traffic [42]. Due to the attack, the system deteriorated to a completely unusable state, once the simulated attack reached a sufficient strength. The lag times increased and the arm movement was very "shaky". Moreover, the behavior changed continuously which prevented the user to learn how to counter these effects. Using the machine learning based detection and mitigation countermeasure described in section II-B, the system remained usable even under the attack, with undetectable lag times and all shakiness removed.

Regarding malicious agents, we identified unexpected possibilities in user authentication. Of course, we had implemented the traditional measures for user authentication such as password protection and physical securing of the facilities. However, VR technology provided us with unique data points with which we could implement more advanced authentication methods. We implemented a simple biometric authentication method by measuring the HMD user height. Much more advanced methods, such as detection of malicious physical interference, could be possible with the VR technology and these are discussed in more detail in the next section.

IV. SECURITY CONSIDERATIONS

Security issues arising around the concept (and even more so, from the various possible implementations of it) are of several types, due to the nature of this idea. First, DTs, like any other proprietary digital content, require protection mechanisms in term of intellectual property: Anyone with sufficient access



to the DT model can theoretically duplicate and distribute it as it might merely be a purely digital asset. Existing methods for securing digital content apply here, to a point. Watermarking [43] and Digital Rights Management (DRM) techniques [44] can protect the intellectual property that is the DT and thus all the company specific software and knowledge that are present in it, but it remains clear from the current state of affairs for audiovisual material, that these solutions can be still easily circumvented [45].

Due to the specific type of digital object that a DT is, one can devise further protection methods (which should be considered only as additions to the state of the art watermarking and DRM systems) for its protection: supposedly, a DT will run on dedicated hardware and a specific set of machines, as opposed to audiovisual material which should be available on as many platforms as possible. Given this restriction, one can imagine means of anchoring the DT to a specific set of hardware components (a server, a desktop computer, a cluster of machines) by the use of a Trusted Platform Module (TPM) [46], for example. By making certain that the software that represents the DT can only run if a successful and unique cryptographic exchange between the software and the hardware is made [47], the DT is bound to this hardware solely. There are numerous research tracks on how to perform this best [48], [49], but the key point is that this is currently achievable using existing technology, as most current machines have such a TPM that can be used. Obviously, the means of implementing this verification in the software/hardware that represents the DT is ad hoc and highly depends on the nature of the twin.

Second, as we have noted in this paper's implementation of a specific type of DT, the need for strong user authentication on all the concerned sites of operation is paramount. In a use case such as this one, state of the art strong authentication techniques, e.g. multi-factor authentication [50], [51], biometric authentication [52], [53], need to be made mandatory in addition to simple physical access restrictions to the facilities where the DT operations are carried out. One only has to consider what can happen in the simple case where a malicious agent gets physical access to the VR control area: it is so simple to disorient and take physical control of someone who is in a VR system, that anybody malicious getting this privileged physical access can wreak havoc with the system. This matter is not only local to the VR site, but becomes also relevant when the other sites, that are part of the operation of the DT, actually perform (physical) actions. In such cases, strong physical and digital authentication practices need to be enforced.

In specific DT use cases, there are some further considerations that need to be taken into account: for the case of remote surgery as in this paper, and furthermore for any kind of sensitive operation, it is important to verify that the operator (i.e. in this context, the human being performing actions within the DT system) is in full capacity. The obvious problematic use case being here that a medical surgeon that has a hand tremor and shaking, should probably not

be authorized to perform. This idea makes the case for a "capabilities check" on top of the strong authentication. This verification can take various forms, depending on the use case, obviously, but in the case of human operators, they are likely to rely on biometric and physical features.

Finally, while the overall system can be fully hardened, security issues can arise due to the supporting network and infrastructure, for example. Again, in the use cases of DTs that help connect and operate two or more sites together, the backbone network supporting such systems can have a critical role. It has been widely reported that network equipment providers have had their equipment tampered with in transport, to embed malware or surveillance features [54]. Without even considering the sort of reach that a malicious network operator would have on DT connection systems (intellectual property theft, impersonation, fake authentication, denial of service...), the problem of having a backdoor inside a supporting network system is that there is a "fox in a henhouse". Any malicious operation becomes to some extent possible, or at least, facilitated: by performing man-in-themiddle attacks, it is possible to intercept and decrypt endto-end encrypted traffic (and depending on the reach of the attacker inside the network, such attacks might go unnoticed for very long periods); it is then possible to duplicate the traffic, including full payloads; it makes the attacker's work easier to plant malware in the network for later data exfiltration and intellectual property theft... In addition, the rise of wireless backbone support for Industry 4.0 and 5G in general, poses another set of potential issues: while most of the network transport in 5G industrial scenarios will be carried out by IP networks, for which vulnerabilities and weaknesses are known (to a point), the idea of local, private, mobile network operators supporting the industrial use case will carry its own set of security matters. The 5G radio access network will provide improved security measures [55], [56] over 4G, but the problem remains that private mobile networks may decide to forego some of these security improvements - for the sake of convenience of implementation, as well as for cost savings reasons, possibly –, leaving the local mobile network open with some security breaches. In this sense, the mobile network part of the overall industrial network infrastructure adds an extra layer of potential vulnerability, if proper measures are not to be taken, when compared to purely wired networks.

These issues (for many of them) are not solely arising in the DT context, but really in most of the future Industry 4.0 paradigm, whereby industrial systems are partially or fully digitized and most systems interconnected via future mobile networks such as 5G. In this regard, many of the current concerns regarding Industry 4.0 security are being actively researched and the best practices that come out should be examined and implemented as much as possible.

V. DISCUSSION

We created a prototype Digital Twin (DT) system which incorporates established internet communication protocols



in achieving a reliable real-time remote control over 4G mobile network connection. However, we needed to implement several pieces of software that actually made the devices and software understand each other. Web APIs (Application Programming Interface) are the standard in developing web applications communicating with different systems. In this case, the nature of the system required establishing of real-time communication between the systems rather than requesting information and waiting for response. For this purpose, the concept of DT is introduced in the ongoing research within the field of Industrial Internet, Industry 4.0 and Internet of Things. The small pieces of software comprise the DT of our prototype system.

Remote replication of human hand operation with viable accuracy was achieved. Based on the test sessions, using the system and controlling a robot accurately is not obvious for most of the people but can be learned rather easily. Some people would need extensive learning, which might be a problem related to the VR technology and not necessarily related to our system implementation.

The behavior of the system was unacceptable with rapid movements. An optimized PI controller for each robot joint could reduce this behavior, but implementing this would need additional research. The major deficit of the system appeared to be the force feedback. A possible improvement for the feedback system would be to incorporate some sort of 3D force feedback device. For example, Mathiassen et al. [57] used a UR5 robotic arm in creating a ultrasound imaging system. In their application, the robot was used to separate the physician from a ultrasound probe in order to prevent musculosceletal disorders and pain. They used a haptic controller device to control the arm. The device is essentially a pen attached to an arm, which measures the movements to be delivered for the controlled robot and simultaneously provides 3D force feedback. In our case, a similar device could be attached to the user to provide 3D force feedback without limiting the users capability to move around.

The feedback system could be further improved by combining visual feedback with the haptic force feedback. As noted, the visual feedback was of low quality and would need better camera equipment and a stereo view. Moreover, the visual feedback modes required the arm to hold the camera and blocked the arm from picking up any other instruments. Having dedicated arms for the camera and instrument would increase flexibility. Since our focus was in the communication part of the system and not in the feedback and usability of the system, we decided to implement only the basic version of feedback.

A. CONCLUSION

In this study, we created the DT little by little and went through the struggle of verifying the functionality. A fundamental promise of DTs is that in the future, systems would be able to establish the communication between devices and verify the level of functionality automatically i.e. without any human writing a single line of code. Examining our prototype DT, the devices should agree that a rapid delivery of packets is required and establish the connection accordingly with UDP or some future protocol suitable for such purpose. Moreover, the systems should solve all the necessary data format, limitation and conversion problems and in the case of failure, should refuse to work and only then request human engineer assistance. After such an incident, similar problems occurring in the future would be automatically solved. This could potentially give birth to a market, which would deal such small pieces of software for specific problems from around the world. Artificial intelligence and machine learning capabilities could eventually be harnessed to tackle such needs. Ultimately, more and more systems would be able to establish communication, understand each other and co-operate for some task without hands on engineering.

In the future, the patient and the surgeon might also possess their own DTs which would enable the system to read historical and real time health and performance data. This leads us to a situation where several DTs are brought into one event for completing a task. Even before the surgeon enters the operation room, she could plan the operation and check the health records of the patient. She might even run some simulations on medication or different strategies for the actual operation. When someone (the patient, a doctor, an insurance company or a health monitoring AI in the patient's personal cloud service) is considering a forthcoming surgery, a digital twin of digital twins needs to be fired up to bring together all the required agents. This means that DTs will be born and terminated for occurring events all the time, which further magnifies the noted security issues.

Investigating mobile networks supported remote surgery bears an ultimate goal of developing a system, which could one day enable such operations to be performed as normal daily activity and not only as ambitious research projects. Our prototype system does not have the capabilities required for that as noticed in the experimental results section. Plenty of development and research in human-machine-interfaces, haptic and visual feedback and network security has to be conducted as those seem to be an crucial part in conducting telesurgery [40]. Some minor medical operations, such as observation and diagnosis, could be conducted with the current capabilities of the system. Moreover, remote medical operations can be simulated and researched with the equipment. Future network technologies, such as 5G, will be required in implementing the high definition video feedback and other more demanding data flows that are necessary for actual remote surgeries.

The development of this prototype system made us painfully aware of the fact that this sort of combination of different engineering fields in novel applications provides plenty of complexity in the development work. Engineers and designers in the fields of robotics, mechatronics, computer sciences, network development and game development each exploit their own tools and methods that are built with their own demands. This is a major obstacle in the realization of the Industry 4.0 and similar concepts and requires research



and consideration in our increasingly cross-disciplinary world.

ACKNOWLEDGMENT

This work was a part of the DIMECC Cyber Trust Program [1].

REFERENCES

- [1] DIMECC. (2017). DIMECC Cyber Trust Program. Accessed: Oct. 22, 2017. [Online]. Available: https://cybertrust.fi/
- [2] S. P. A. Datta. (2016). "Emergence of digital twins." [Online]. Available: https://arxiv.org/abs/1610.06467
- [3] J. Autiosalo, "Platform for industrial Internet and digital twin focused education, research, and innovation: Ilmatar the overhead crane," in Proc. IEEE 4th World Forum Internet Things (WF-IoT), Feb. 2018, pp. 241–244.
- [4] E. Glaessgen and D. Stargel, "The digital twin paradigm for future NASA and U.S. air force vehicles," in *Proc. 53rd AIAA/ASME/ASCE/AHS/ASC Struct.*, Struct. Dyn. Mater. Conf., Honolulu, HI, USA, Apr. 2012. doi: 10.2514/6.2012-1818.
- [5] Q. Zhang, J. Liu, and G. Zhao. (2018). "Towards 5G enabled tactile robotic telesurgery." [Online]. Available: https://arxiv.org/abs/1803.03586
- [6] T. L. McDonald, "Are we having virtual fun yet?" PC Gamer US, vol. 1, no. 3, pp. 44–49, Aug. 1994. [Online]. Available: https://imgur.com/gallery/7G1h9/
- [7] R. Hollander. (Aug. 2017). Here's Why AR and VR are set to Take Off. Accessed: Jan. 5, 2018. [Online]. Available: http://www.businessinsider.com/ar-vr-2017-8?r=US&IR=T
- [8] O. Holmes, M. S. Banks, and H. Farid, "Assessing and improving the identification of computer-generated portraits," ACM Trans. Appl. Percept., vol. 13, no. 2, 2016, Art. no. 7. doi: 10.1145/2871714.
- [9] H. Laaki, K. Kaurila, K. Ots, V. Nuckchady, and P. Belimpasakis, "Augmenting virtual worlds with real-life data from mobile devices," in *Proc. IEEE Virtual Reality Conf. (VR)*, Mar. 2010, pp. 281–282.
- [10] Alexandra Thompson. There's a Patient Under There Somewhere: The Incredible Medical Robots That Have Saved a Record Number of Men From Deadly Prostate Cancer.

 Accessed: Dec. 10, 2018. [Online]. Available: http://www.dailymail.co.uk/health/article-5114551/Robot-surgery-save-men-prostate-cancer.html
- [11] RoboticOncology.com. Benefits of Robotic Prostate Cancer Surgery. Accessed: Jan. 5, 2018. [Online]. Available: http://www.roboticoncology.com/robotic-prostate-surgery/
- [12] J. Marescaux et al., "Transatlantic robot-assisted telesurgery," Nature, vol. 413, no. 6854, pp. 379–380, 2001.
- [13] Y. Miao, Y. Jiang, L. Peng, M. S. Hossain, and G. Muhammad, "Telesurgery robot based on 5G tactile Internet," *Mobile Netw. Appl.*, vol. 23, no. 6, pp. 1645–1654, Dec. 2018. doi: 10.1007/s11036-018-1110-3.
- [14] The Matrix, Wachowski Brothers, USA, 1999.
- [15] Bethesda Game Studios. (Dec. 2017). Fallout 4 VR. [Online]. Available: http://store.steampowered.com/app/611660/Fallout_4_VR/
- [16] Enscape GmbH. Enscape3D. Accessed: Nov. 7, 2017. [Online]. Available: https://enscape3d.com/
- [17] Google Inc. Tilt Brush. Accessed: Nov. 7, 2017. [Online]. Available: https://www.tiltbrush.com/
- [18] Virtalis Limited. VR4CAD. Accessed: Nov. 7, 2017. [Online]. Available: https://www.virtalis.com/yr4cad/
- [19] A. G. Iñárritu, "Flesh and Sand (Spanish: Carne y arena)," Legendary Pictures, USA, Tech. Rep., May 2017.
- [20] HTC Corporation. ViveâĎĆ Room-Scale VR. Accessed: Jan. 6, 2018. [Online]. Available: https://www.vive.com/eu/product/
- [21] J. Velasco. (Sep. 2016). HTC Vive Review. Accessed: Jan. 6, 2018. [Online]. Available: https://vrsource.com/htc-vive-review-3443/
- [22] Oculus VR. Oculus Rift. Accessed: Jan. 8, 2018. [Online]. Available: https://www.oculus.com/rift/
- [23] O. Kreylos. (May 2016). Lighthouse Tracking Examined. Accessed: Sep. 13, 2017. [Online]. Available: http://doc-ok.org/?p=1478
- [24] S. Niku, Introduction to Robotics: Mechanics and Control, 2nd ed. Hoboken, NJ, USA: Wiley, 2010.
- [25] Tesla. (Jun. 2012). Inside Tesla 06.12.12—Paint. Accessed: Jan. 12, 2018. [Online]. Available: https://youtu.be/vXntSUtoEFc

- [26] Robots and Robotic Devices—Safety Requirements for Industrial Robots— Part 1: Robots, Standard ISO 10218-1:2011, ISO, Geneva, Switzerland, 2011
- [27] Robots and Robotic Devices—Collaborative Robots, Standard ISO/TS 15066:2016, ISO, Geneva, Switzerland, 2016.
- [28] Universal Robots A/S. (Sep. 29, 2017). The URScript Programming Language. [Online]. Available: https://www.universal-robots.com/download/?option=22197#section21968
- [29] Universal Robots A/S. (Sep. 29, 2017). User Manual UR3/CB3. [Online]. Available: https://www.universal-robots.com/download/?option=21983#section21970
- [30] Universal Robots A/S. (Jan. 5, 2018). UR3 Robot. [Online]. Available: https://www.universal-robots.com/products/ur3-robot/
- [31] Unity Technologies. (Jan. 6, 2018). Unity. [Online]. Available: https://unity3d.com/unity
- [32] ValveSoftware. (Jan. 7, 2018). OpenVR. [Online]. Available: https://github.com/ValveSoftware/openvr/wiki/API-Documentation
- [33] R. Braden, Requirements for Internet Hosts—Communication Layers, document RFC 1122, Internet Engineering Task Force, Oct. 1989. [Online]. Available: http://tools.ietf.org/html/rfc1122
- [34] O. Ravn, N. Andersen, and T. Andersen. UR10 Performance Analysis. [Online]. Available: http://orbit.dtu.dk/ws/files/105275650/ ur10_performance_analysis.pdf
- [35] H. Laaki, "Virtual reality control system for industrial robot," M.S. thesis, Aalto Univ. School Eng., Helsinki, Finland, 2018. [Online]. Available: http://urn.fi/URN:NBN:fi:aalto-201804031991
- [36] A. Kalliola *et al.*, "Learning flow characteristics distributions with elm for distributed denial of service detection and mitigation," in *Proc. ELM*, J. Cao, E. Cambria, A. Lendasse, Y. Miche, and C. M. Vong, Eds. Cham, Switzerland: Springer, 2018, pp. 129–143.
- [37] Heikki Laaki. (Jan. 2018). UR3 VR Remote Control: Tracking Demo. Accessed: Jan. 8, 2018. [Online]. Available:https://youtu.be/Sfv-Mot8cvo
- [38] Venimis Com. (Sep. 2017). Cybertrust Seminar-Yoan Miche-Nokia Bell Labs. Accessed: Jan. 17, 2018. [Online]. Available: https://youtu.be/k9bYKdxGVDA
- [39] S. Sharples, S. Cobb, A. Moody, and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): Comparison of head mounted display (HMD), desktop and projection display systems," *Displays*, vol. 29, no. 2, pp. 58–69, 2008. [Online]. Available: http://www. sciencedirect.com/science/article/pii/S014193820700100X
- [40] G. Tholey, J. P. Desai, and A. E. Castellanos, "Force feedback plays a significant role in minimally invasive surgery: Results and analysis," *Ann. Surg.*, vol. 241, no. 1, pp. 102–109, 2005.
- [41] L. Mearian. (Jan. 5, 2018). Hospital Tests Lag Time for Robotic Surgery 1,200 Miles Away From Doctor. Accessed: Nov. 2009. [Online]. Available: https://www.computerworld.com/article/2927471/healthcare-it/robot-performs-test-surgery-1200-miles-away-from-doctor.html
- [42] United States Computer Emergency Readiness Team. (Jan. 5, 2018). Security Tip (ST04-015). [Online]. Available: https://www.us-cert.gov/ncas/tips/ST04-015
- [43] S. Katzenbeisser and F. A. P. Petitcolas, Information Hiding Techniques for Steganography and Digital Watermarking. Norwood, MA, USA: Artech House. 2000.
- [44] S. R. Subramanya and B. K. Yi, "Digital rights management," *IEEE Potentials*, vol. 25, no. 2, pp. 31–34, Mar. 2006.
- [45] B. Widder. (Feb. 2015). DRM Getting You Down? Here's How to Strip Your Music and Movies of Restrictions. [Online]. Available: https://www.digitaltrends.com/home-theater/how-to-remove-drm-from-music-and-movie-files/
- [46] T. Morris, "Trusted platform module," in *Encyclopedia of Cryptography and Security*. Boston, MA, USA: Springer, 2011, pp. 1332–1335.
- [47] W. Arthur, D. Challener, and K. Goldman, "Quick tutorial on TPM 2.0," in A Practical Guide to TPM 2.0. Berkeley, CA, USA: Springer, 2015, pp. 23–37.
- [48] C. S. Nachenberg and B. McCorkendale, "Securing executable content using a trusted computing platform," U.S. Patent 7 694 139, Apr. 6, 2010.
- [49] V. R. Scarlata and C. V. Rozas, "Method and apparatus for providing secure virtualization of a trusted platform module," U.S. Patent 7 590 867, Sep. 15, 2009.
- [50] V. Kirillin, S. Zemlyanskiy, D. A. Baranov, and V. Podoshvin, "Multi factor user authentication," U.S. Patent 13 464 504, Nov. 7, 2013.
- [51] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," *Proc. IEEE*, vol. 91, no. 12, pp. 2021–2040, Dec. 2003.



- [52] J. Wayman, A. Jain, D. Maltoni, and D. Maio, "An introduction to biometric authentication systems," in *Biometric Systems*. London, U.K.: Springer, 2005, pp. 1–20.
- [53] P. Tuyls, A. H. M. Akkermans, T. A. M. Kevenaar, G.-J. Schrijen, A. M. Bazen, and R. N. J. Veldhuis, "Practical biometric authentication with template protection," in *Proc. Int. Conf. Audio- Video-Based Biomet*ric Person Authentication. Berlin, Germany: Springer, 2005, pp. 436–446.
- [54] S. Gallagher. (May 2014). Photos of an NSA 'Upgrade' Factory Show Cisco Router Getting Implant. [Online]. Available: https://arstechnica.com/tech-policy/2014/05/photos-of-an-nsa-upgrade-factory-show-cisco-router-getting-implant/
- [55] D. Fang, Y. Qian, and R. Q. Hu, "Security for 5G mobile wireless networks," *IEEE Access*, vol. 6, pp. 4850–4874, 2018.
- [56] 3GPP 5G Security. Accessed: Dec. 10, 2018. [Online]. Available: http://www.3gpp.org/news-events/3gpp-news/1975-sec_5g
- [57] K. Mathiassen, J. E. Fjellin, K. Glette, P. K. Hol, and O. J. Elle, "An ultrasound robotic system using the commercial robot ur5," *Frontiers Robot. AI*, vol. 3, p. 1, Feb. 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/frobt.2016.00001



HEIKKI LAAKI was born in Toijala, Finland, in 1992. He received the B.S. and M.S. degrees in mechanical engineering from Aalto University, Espoo, Finland, in 2016 and 2018, respectively, where he is currently pursuing the Ph.D. degree.

In 2017, he was a Thesis Worker at the Cybersecurity Research Group, Nokia Bell Labs, and wrote his thesis on the title Virtual Reality Control System for Industrial Robot. He served in the military in Finnish Air Force on the Pilot Reserve

Officer Course, from 2011 to 2012. His research interests include virtual reality and augmented reality, human machine interfaces, and digital twins.



YOAN MICHE was born in France, in 1983. He received the Engineer degree from the TELE-COM, Institut National Polytechnique de Grenoble (INPG), France, in 2006, the master's degree in signal, image, speak and telecom from ENSERG, INPG, in 2006, and the D.Sc. degree from INPG and Aalto University, Finland. He worked in the Information and Computer Science Lab, Aalto University, as a postdoctoral position for four years. He is currently the Head of cybersecurity research with Nokia Bell Labs, Finland.



KARI TAMMI (M'15) was born in 1974. He received the M.Sc., Lic.Sc., and D.Sc. degrees from the Helsinki University of Technology, in 1999, 2003, and 2007, respectively. He received a Teacher's pedagogical qualification from the Häme University of Applied Sciences, in 2017. He was a Researcher with CERN, the European Organization for Nuclear Research, from 1997 to 2000, and a Postdoctoral Researcher with North Carolina State University, USA, from 2007 to

2008. From 2000 to 2015, he was a Research Professor, a Research Manager, the Team Leader, and other positions at the VTT Technical Research Centre, Finland. He has been an Associate Professor with Aalto University, since 2015. He currently serves in the Finnish Administrative Supreme Court as a Chief Engineer Counselor. He has authored more than 60 peer-reviewed publications cited in over 1500 other publications. He is a member of the Finnish Academy of Technology. He serves as the Deputy Chair for IFTOMM Finland.

. . .